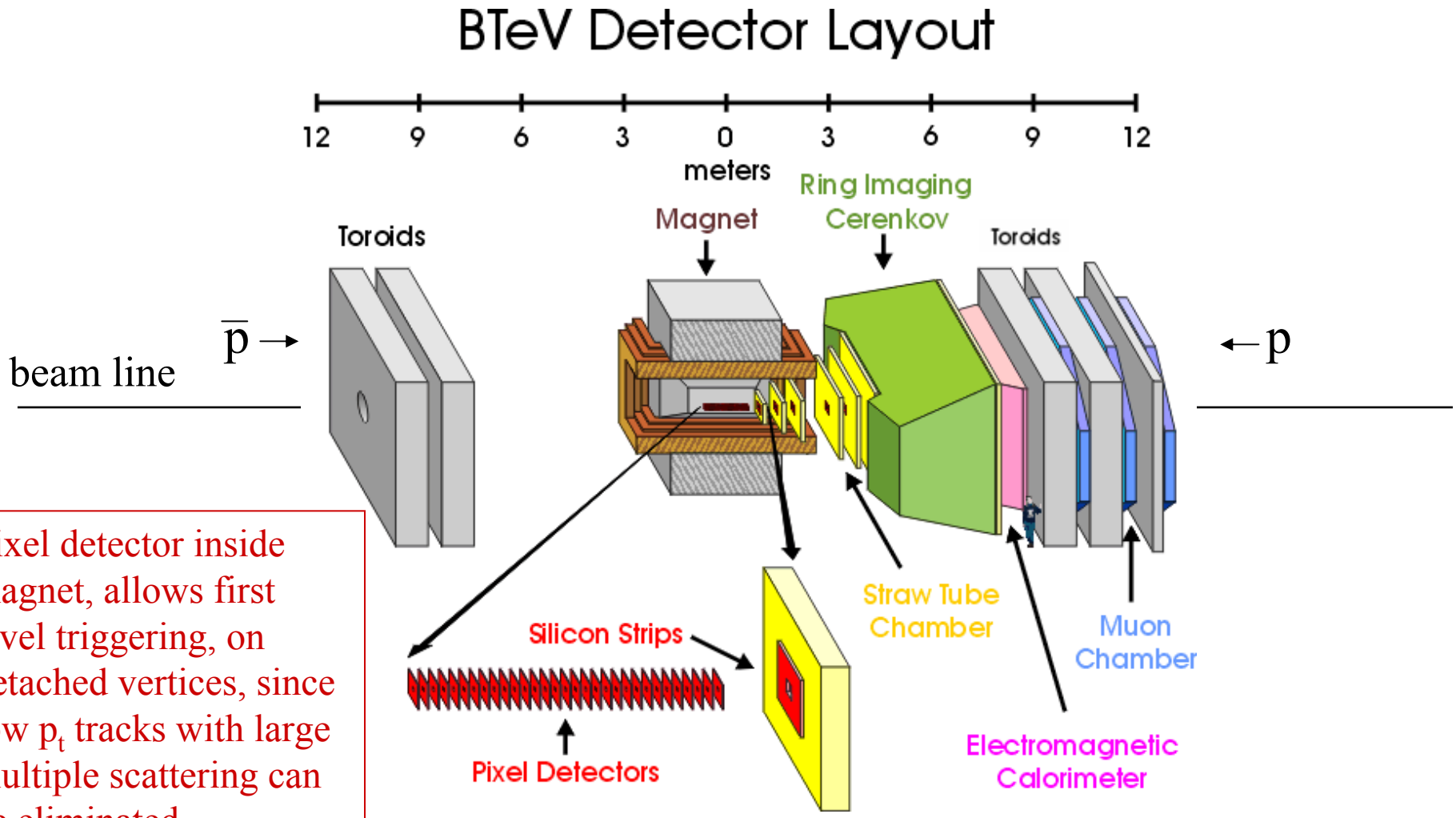
BTeV Trigger/DAQ Innovations

Margaret Votava
Fermilab
On Behalf of BTeV Trigger and DAQ Groups

BTeV Background

- B-physics collider experiment proposed at Fermilab
  - http://www-btev.fnal.gov
  - 170 physicists from 30 universities/institutions worldwide
  - $193M construction project ($141M base + $52M contingency)
    - DAQ and Trigger were 2 separate subprojects accounting for 17% of the project cost combined.
- Timeline
  - In R&D since 1996!
  - Construction to have begun in FY2005. Complete in FY2010.
  - Successfully passed Department of Energy CD2/CD3a review (*i.e.*, limited construction funds available) in December 2004.
  - Cancelled abruptly in February, 2005 by the Department of Energy because of funding constraints.

The BTeV Detector



Pixel detector inside magnet, allows first level triggering, on detached vertices, since low p_t tracks with large multiple scattering can be eliminated

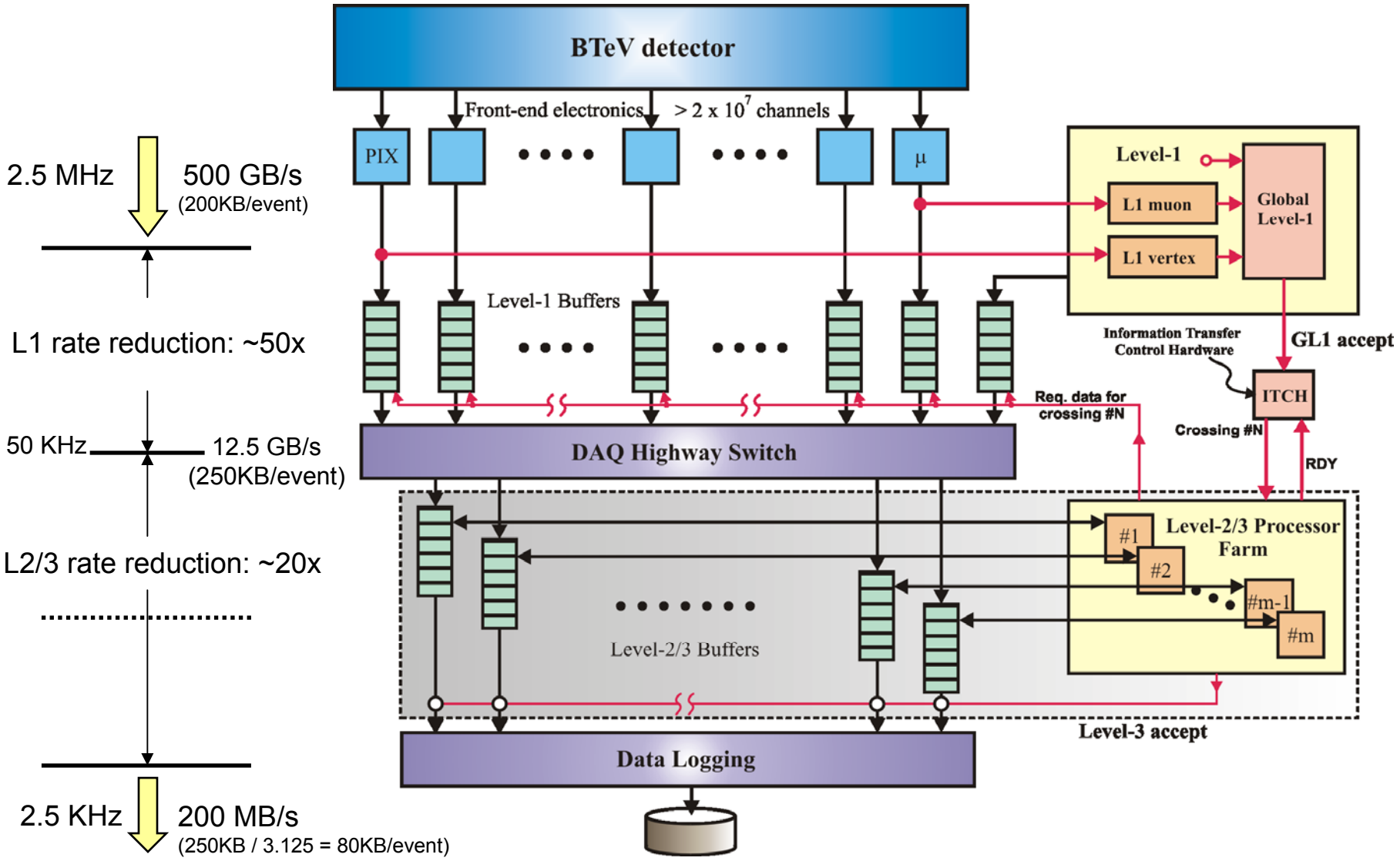
Online Requirements

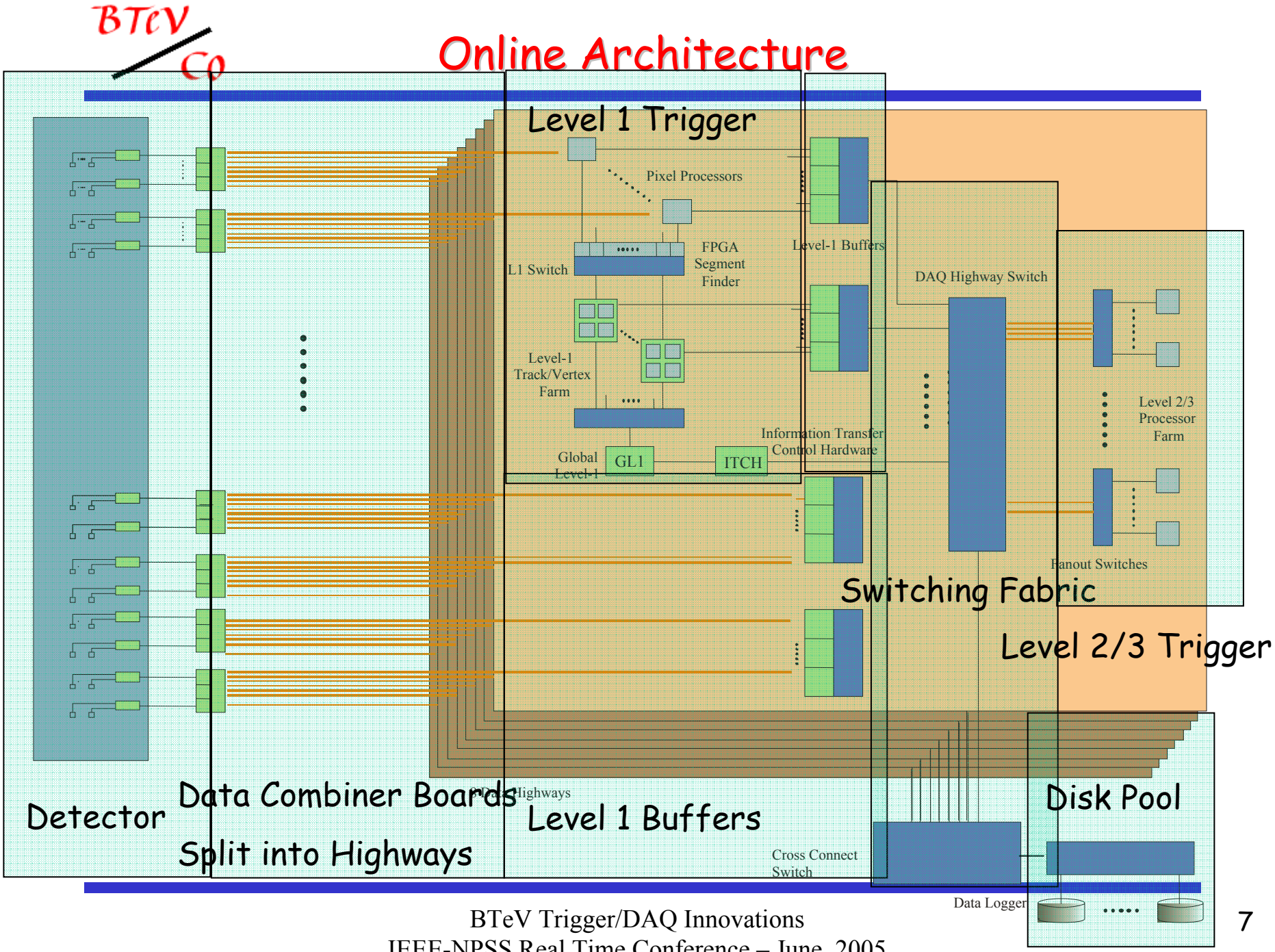
- The challenge for the BTeV trigger and data acquisition system is to reconstruct particle tracks and interaction vertices for EVERY interaction that occurs in the BTeV detector, and to select interactions with B decays.
- Reject > 99.9% of background. Keep > 50% of B events.
  - The trigger performs this task using 3 levels, referred to as Levels 1, 2, and 3:
    - "L1" – looks at every interaction and rejects at least 98% of min. bias background
    - "L2" – uses L1 computed results & performs more refined analyses for data selection
    - "L3" – rejects additional background and performs data-quality monitoring
  - The data acquisition system saves all of the data in memory for as long as necessary to analyze each interaction, and moves data to L2/3 processing units and archival data storage for selected interactions.
    - Complex algorithms => long latencies (on the order of 1 msec for L1)
- The key ingredients that make it possible to meet this challenge:
  - BTeV pixel detector with its exceptional pattern recognition capabilities
  - Rapid development in technology – FPGAs, commodity processors, networking
- I/O rates
  - 24 million channels, dominated by Pixel detector
  - Input rate of 500 Gbytes/sec (2.5MHz x 200KBytes/event)
  - Output rate of 200 Mbytese/sec (2.5 KHz x 80 Kbytes/event)
    - 1 Petabyte/year

Value Engineering Approach

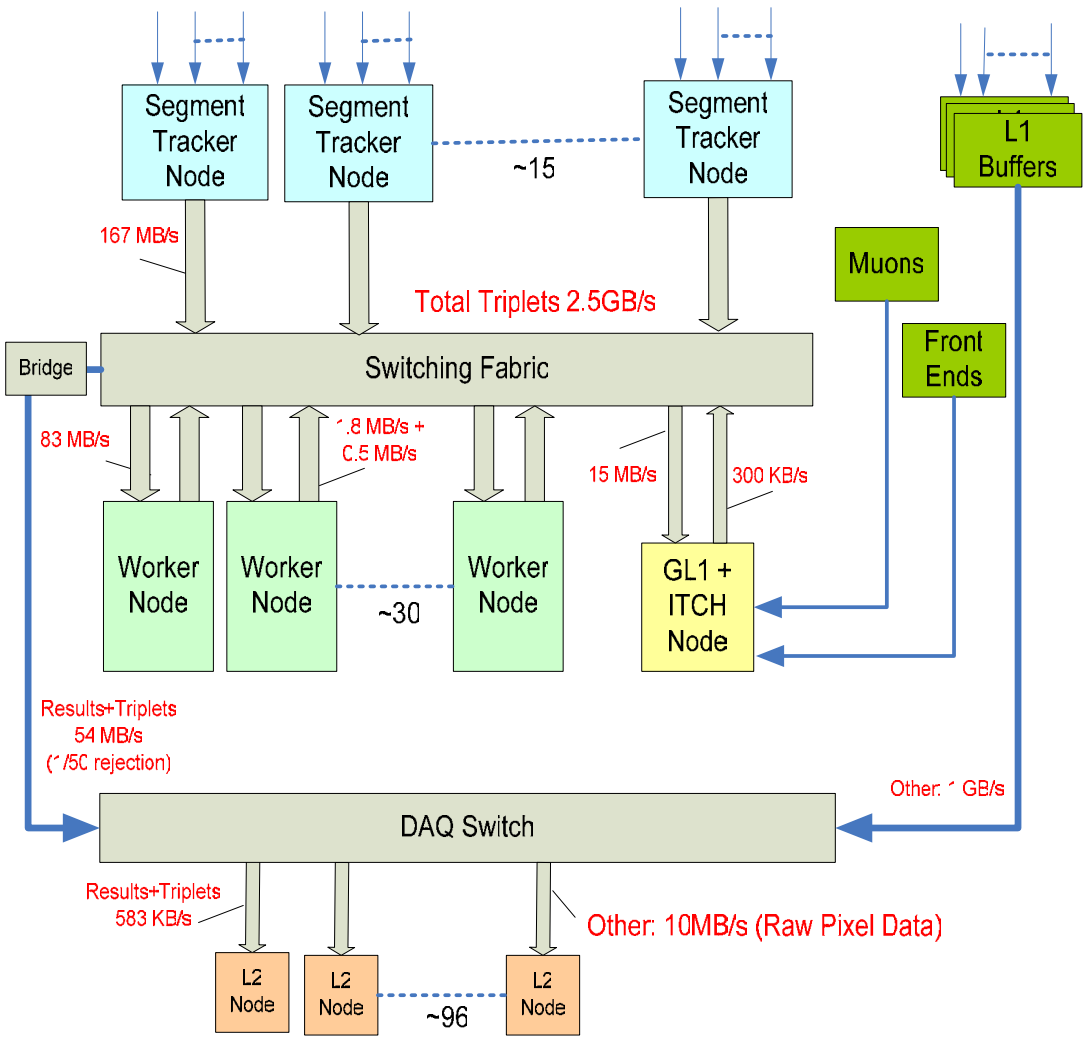
- Fast data links between detector and online system
- Sophisticated L1 tracking trigger
- Commodity hardware wherever possible.
- Use inexpensive DRAM-type buffer memory off detector
- Only point to point serial links (copper and optical)
- Subdivide system into 8 parallel and independent data streams (a.k.a., highways) each operating with 1/8th of the overall data.

Block Diagram of Trigger & DAQ Data Flow




BTeV
Co
Online Architecture
Level 1 Trigger
Pixel Processors
L1 Switch
FPGA
Segment
Finder
Level-1 Buffers
Level-1
Track/Vertex
Farm
DAQ Highway Switch
Level 2/3
Processor
Farm
Information Transfer
Control Hardware
Global
Level-1
GL1
ITCH
Fanout Switches
Switching Fabric
Level 2/3 Trigger
Detector
Data Combiner Boards
Split into Highways
Level 1 Buffers
Cross Connect
Switch
Disk Pool
Data Logger

L1 Pixel Trigger Architecture



- Data driven receiver must accept what is sent (no handshaking)
- Parallel operation limits effect of hardware failure
- Asynchronous
- Decision latency and event size variable and uncorrelated
- Deadtimeless
- Scalable by modularity of design
- More detailed information in other talks:
  - The Application of Tiny Triplet Finder (TTF) in BTeV Pixel Trigger (S6-3)
  - Integrated Upstream Parasitic Event Building Architecture for BTeV Level 1 Pixel Trigger System (P8-3)
  - A Commodity Solution Based High Data Rate Asynchronous Trigger System for Hadron Collider (S14-1)

RTES Project

- Real Time Embedded Systems
  - Collaboration of physicists and computer scientists
    - Funded by the National Science Foundation (NSF)
  - Fault tolerant, adaptive solutions
  - RT2003 talk by L. Piccoli – demonstrated ARMORs (one component) for process management
  - http://www-btev.fnal.gov/public/hep/detector/rtes/ - RTES home page
- To be presented tomorrow morning (S9-1)
  - Mike Haney (UofI) – BTeV and Beyond

Partitioning

- Partitioning is the logical concept of running multiple data acquisition sessions in parallel.
  - Not to be confused with multiple highways, which is a physical architecture of the data path.
    - A single partition can be collecting data from 1-8 highways and/or
    - Multiple partitions can be running on any particular highway
  - Myriad of implementation possibilities given the 8 highway architecture and run time configurable routing tables in the data combiner boards.
- Partitions support commissioning
  - Standalone test of subdetectors and electronics
  - Test subdetectors and electronics in gangs
  - Data collection could range from a very low rate (a few events per second) to full rate
- Partitions support operations
  - Obey prime directive => collect maximum amount of physics events.
  - Corral spare cycles on the online trigger farm for offline processing in a non-disruptive way.
  - Test new trigger algorithms

Partitioning Model in BTeV

- Human run coordinator will:
  - Establish number of working highways
  - Coordinate data taking runs during this period (e.g., knowledge of who will need pixel trigger)
  - Start trigger
- During commissioning
  - Each Data taker (*i.e.,* partition owner) will:
    - Select data sources (in units of L1 buffers) for read/write or read only
    - Select data sinks (in units of 12 worker nodes)
    - Select set of trigger tables
- During operations
  - Worker nodes can be manually assigned to offline partition (for long down times) in units of regional managers
  - Worker nodes can automatically shift to offline partition as luminosity decreases (influenced by RTES)

Highlights

- Credible Design – favorable comments by reviewers
- Standardized inputs to Trigger/DAQ as early as possible. Data Combiner Boards were the single entry point to the online system
- Highway architecture
  - Individual highways operating at 1/8 of full data rate
  - Control overhead more manageable
  - Network bandwidth used more efficiently (larger packet sizes)
- Asynchronous L1 pixel trigger on commodity hardware. Compared to custom solutions the commodity choice:
  - Has lower cost
  - Has lower risk
  - Requires an increase in power & cooling
  - Is easier to build (less custom hardware to design)
  - Requires less labor (less engineering & no DSP programming)
  - Has an easier upgrade path (easier to add or replace processors)
- Reliable, fault adaptive system
- Offline capabilities on online farm

Contacts

- Contact us:
  - Margaret Votava: votava@fnal.gov
  - Erik Gottschalk: erik@fnal.gov
- Talk to us at the conference:
  - Jin Yuan Wu
    - The Application of Tiny Triplet Finder (TTF) in BTeV Pixel Trigger (S6-3)
    - Integrated Upstream Parasitic Event Building Architecture for BTeV Level 1 Pixel Trigger System (P8-3)
  - Mike Haney
    - The RTES Project - BTeV, and Beyond (S9-1)
  - Mike Wang (disguised in Jin Yuan)
    - A Commodity Solution Based High Data Rate Asynchronous Trigger System for Hadron Collider (S14-1)
  - Luciano Piccoli
- Surf the web:
  - http://www-btev.fnal.gov – BTeV home page include links to TDR
  - http://www-btev.fnal.gov/public/hep/detector/rtes/ - RTES home page

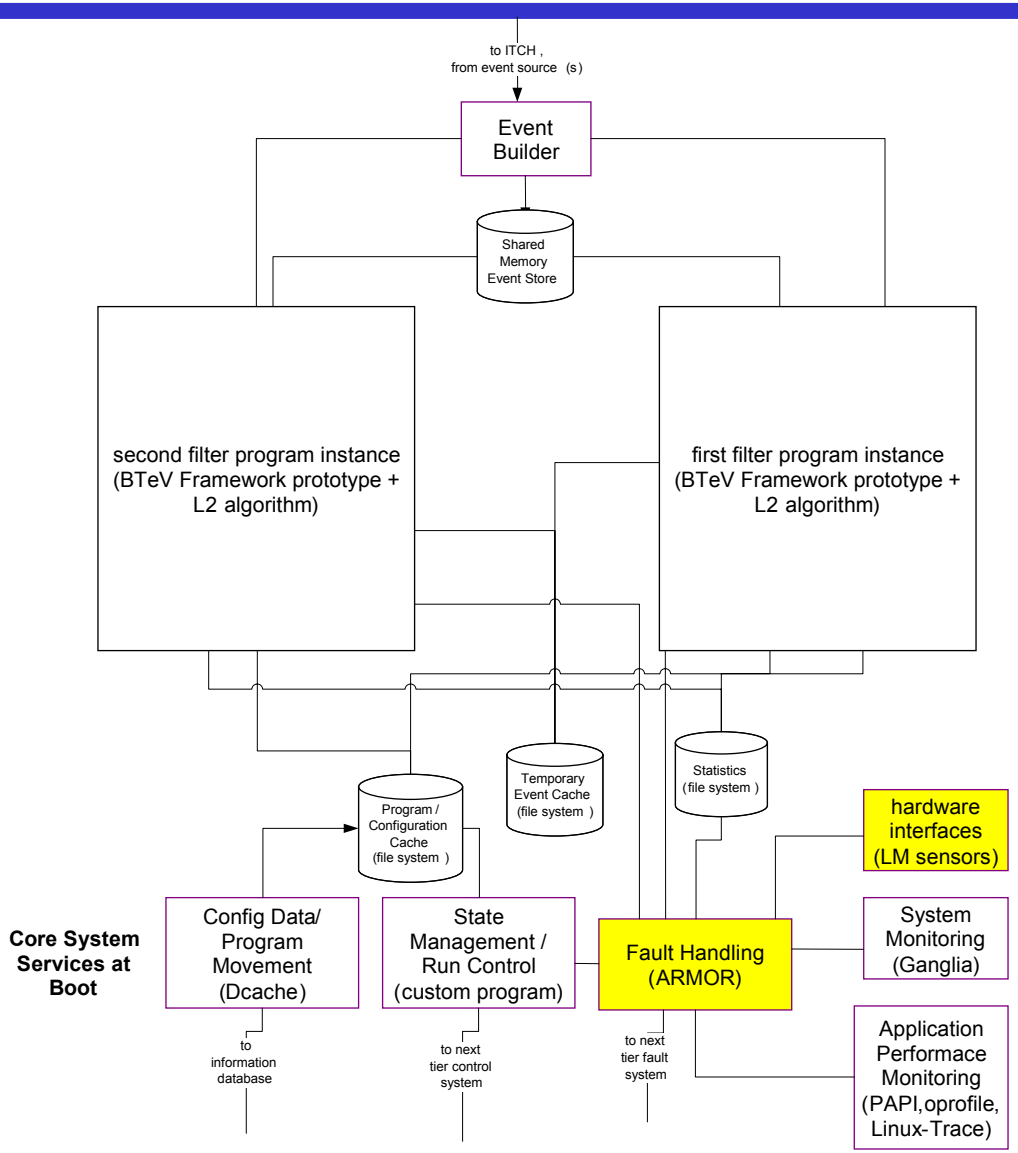
BACKUP SLIDES

SLIDES THAT FOLLOW HAVE ADDITIONAL INFORMATION

Online Architecture Credible

- Detailed Work Breakdown Structure (WBS) for complete experiment.
  - Fully burdened resource loaded schedule
  - ~6000 activities for trigger/DAQ.
  - Risk analysis understood and mitigation strategies planned.
  - Included both cost and schedule contingency. Online system completed at least 9 months before it was needed.
- Heavily reviewed.
  - Went before P5 subpanel (twice!)
  - Department of Energy reviews CD0, CD1, and CD2/3a, each preceded by an internal Director's review.
  - BTeV passed the final review – CD2/3a with notably favorable comments from the review committee.
- Prototyping
  - L1 Trigger prototyped
  - Prepilot being assembled

Worker Node Software Architecture



Trigger/DAQ Glossary

DCB	Data Combiner Board
FPGA	Field Programmable Gate Array
GBE	Gigabit Ethernet
GL1	Global Level 1
HIGHWAY	One of eight segments of the data path
ITCH	Information Transfer Control Hardware
L1B	Level 1 Buffer
Partition	Logical concept to define a data acquisition run
PTSM	Pixel Trigger Supervisor and Monitor
RTES	Real-Time Embedded Systems